The genetic basis of amyotrophic lateral sclerosis: recent breakthroughs

Abstract: Deciphering the genetic architecture of amyotrophic lateral sclerosis (ALS), an adult-onset neurodegenerative disorder of the motor neuron system, is important to understand the etiology of this fatal disease as well as to develop customized ALS therapies based on the patient’s genetic fingerprint. In this review, we discuss the genetic basis of ALS, and attempt to link the causal genes to three highly interrelated pathogenic mechanisms: dysproteostasis, RNA dysregulation, and axon dysfunction. In addition, we address the clinical and biological implications of these genetic findings. Furthermore, we explore to what extent genetic knowledge can be converted into targeted and personalized treatments.

Keywords: amyotrophic lateral sclerosis, frontotemporal dementia, genetics, disease modifiers, personalized medicine

Introduction

Amyotrophic lateral sclerosis (ALS)

ALS is a severely disabling and lethal neurodegenerative disorder that affects ~2–3 per 100,000 individuals worldwide. It generally strikes people over the age of 50, although juvenile ALS occurs earlier. It is usually fatal within 3–5 years after symptom onset. Selective dysfunction and dying of motor neurons in the spinal cord, brainstem, and motor cortex are the key features of ALS. Consequently, the major symptoms consist of muscle weakness and paralysis, which emerge when motor axonal retraction occurs. Cognitive dysfunction is an additional finding in a subset of patients. The majority of patients (~90%) are classified as being sporadic ALS (SALS) because they are not aware of other affected family members. Only in 10% of ALS patients a family history of disease is evident; that form is referred to as familial ALS (FALS). The borders between FALS and SALS are gradually fading, as discussed in the “FALS and SALS” section. Despite the progress made in ALS research so far, current clinical practice fails to halt or even slow down the disease course significantly. Riluzole, the only FDA-approved drug for ALS, results in a survival benefit, which is however limited and not noticeable for patients.

Disease variability

Although the clinical picture of ALS may seem remarkably stereotypic, phenotypic and genetic heterogeneity is a persistent feature of the disease. Age of onset, clinical manifestation, type of motor neuron involvement, disease duration, survival, and many other disease parameters vary significantly.

Cognitive dysfunction is evident in some patients. Recent evidence indeed shows ALS to be closely related to frontotemporal dementia (FTD). FTD is a degenerative...
disease affecting large neurons in the frontal and temporal cortex. It usually manifests itself through behavioral or language abnormalities. Approximately 40% of FTD patients have a family history of the disorder, progranulin (PGRN) and MAPT mutations being the most frequently encountered causes for FTD until recently. Many ALS patients have neuropsychological (usually behavioral) abnormalities suggestive for FTD, a minority of them meeting the criteria for the additional diagnosis of FTD. The mirror finding also holds true: many FTD patients have motor abnormalities of variable severity (mostly amyotrophy). Therefore, ALS and FTD are thought to represent the extremes of one disease spectrum, called the ALS/FTD spectrum.

In addition to the marked phenotypic variability, there also is marked genetic variability. The number of genes in which pathogenic mutations cause ALS is steadily increasing (Table 1). Hitherto, more than 25 ALS-related genes have been identified, explaining the cause of more than half of FALS. Thus, mutations in, at first glance, very different genes induce a very similar (or even identical) clinical picture, but the same mutation may induce quite variable phenotypes both in terms of onset and survival, or type of involvement. This generated an interesting but rather semantic debate on whether ALS is a true disease entity, or rather embodies a heterogeneous group of diseases with motor neuron degeneration as a common characteristic.

This progress in genetics has allowed the generation of disease models, which, hopefully, will serve as tools for the rational design and testing of therapeutic approaches. Furthermore, the boom in genetic discoveries has led to the notion that therapeutic interventions will need to be personalized, with ALS patients receiving the most suited therapy based on their genetic fingerprint.

FALS and SALS

Whether an individual patient has SALS or FALS is clinically decided based upon careful analysis of family history. However, mutations in typical ALS genes are found in a sizeable portion of patients labeled as sporadic. Incomplete penetrance, insufficient knowledge of family history, and diagnostic errors related to family members’ are, among others, reasons why the genetic predisposition is unnoticed. This has profound implications for clinicians in terms of correctly informing the family. Therefore, many ALS specialists consider genetic testing in all patients independent of a family history of disease. In addition, the finding that SALS patients actually have FALS, and the notion that some mutations have limited penetrance and thus may be considered as risk factors, have led to the hypothesis that all ALS is inheritable.

A genetic basis for ALS without a familial context has indeed been addressed in twin studies and genome-wide association studies (GWAS), and is now being explored using recently developed sequencing techniques.

In this review, we first discuss the genetic factors implicated in ALS and classify them in three possible disease mechanisms: protein homeostasis, RNA metabolism, and cytoskeletal integrity (Figure 1A–C). This classification is somewhat arbitrary and premature, as the exact mechanisms through which the mutations discussed induce motor neuron degeneration are still incompletely understood. Second, we assess the clinical and biological implications of neurogenetics and GWAS, and dig into genetic factors that act as disease modifiers. Third, we discuss what new genetic approaches may mean for the understanding of ALS. As the discovery of ALS-causing and modifying genes yields possible targets for intervention in cellular pathways, we critically look at the potential of translating this knowledge into target-based therapeutics. We conclude this review by looking into the possibility of customized ALS therapy based on genetic information.

ALS: collapse of the proteostasis machinery

A common hallmark of several neurodegenerative disorders is the accumulation of proteinaceous deposits. According to current thinking, accumulation of mutant, misfolding-prone proteins in vulnerable neurons and neighboring nonneuronal cells is induced by cellular stressors, of which one may be aging. By interfering with the proteostatic machinery, these protein aggregates fuel a self-reinforcing vicious circle of misfolded protein buildup and subsequent disturbance of various cellular processes. The convergence of the resulting effects within multiple cell types ultimately results in motor axonal retraction and motor neuronal death. There is ample experimental support for protein misfolding, protein accumulation, and proteostatic machinery dysfunction in ALS. However, the temporal relationship between these elements, their pathogenic contribution, the identity and role of cellular stresses, and the primary significance of these factors have been very difficult to firmly establish.

Superoxide dismutase 1 (SOD1): the old-timer among causal ALS genes

The groundbreaking discovery of the first ALS gene, SOD1, more than 20 years ago, turned the attention to hereditary ALS. Mutations in this gene account for ~12% of FALS and are found in ~1% of SALS. SOD1 is a ubiquitously expressed enzyme that catalyzes the dismutation of
Table 1 List of selected genes associated with ALS and FTD

<table>
<thead>
<tr>
<th>Genes</th>
<th>Locus</th>
<th>Inheritance</th>
<th>Discovery method</th>
<th>Suggested role in ALS*</th>
<th>Involvement in the ALS-FTD spectrum</th>
<th>Putative protein function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Causal ALS genes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PGRN</td>
<td>17q21.31</td>
<td>AD</td>
<td>Candidate gene, association study</td>
<td>Established</td>
<td>FTD</td>
<td>Growth factor</td>
<td>12–14</td>
</tr>
<tr>
<td>MAPT</td>
<td>17q21.31</td>
<td>NA</td>
<td>Candidate gene, association study</td>
<td>Established</td>
<td>FTD</td>
<td>Cytoskeleton</td>
<td>15</td>
</tr>
<tr>
<td>SOD1</td>
<td>21q22 (AR rare)</td>
<td>AD</td>
<td>Linkage</td>
<td>Established</td>
<td>ALS</td>
<td>Superoxide metabolism, toxic aggregation</td>
<td>31</td>
</tr>
<tr>
<td>UBQLN2</td>
<td>Xp11</td>
<td>XL</td>
<td>Linkage</td>
<td>Established</td>
<td>ALS, ALS/FTD</td>
<td>Proteasome</td>
<td>45</td>
</tr>
<tr>
<td>SQSTM1</td>
<td>5q35</td>
<td>AD</td>
<td>Candidate gene</td>
<td>Established</td>
<td>ALS, ALS/FTD</td>
<td>Ubiquitination, autophagy</td>
<td>55</td>
</tr>
<tr>
<td>VCP</td>
<td>9p13</td>
<td>AD</td>
<td>Candidate gene</td>
<td>Established</td>
<td>ALS</td>
<td>Proteasome, vesicle trafficking</td>
<td>56</td>
</tr>
<tr>
<td>OPTN</td>
<td>10p13</td>
<td>AD</td>
<td>Homozygosity mapping</td>
<td>Established</td>
<td>ALS</td>
<td>Autophagy</td>
<td>63</td>
</tr>
<tr>
<td>CHMP2B</td>
<td>3p11.2</td>
<td>AD</td>
<td>Linkage, candidate gene</td>
<td>Established</td>
<td>ALS, FTD (rare)</td>
<td>Vesicle trafficking</td>
<td>68</td>
</tr>
<tr>
<td>TARDBP</td>
<td>1p36</td>
<td>AD</td>
<td>Linkage, candidate gene</td>
<td>Established</td>
<td>ALS (ALS/FTD)</td>
<td>RNA metabolism</td>
<td>71</td>
</tr>
<tr>
<td>FUS</td>
<td>16p11</td>
<td>AD</td>
<td>Linkage, candidate gene</td>
<td>Established</td>
<td>ALS</td>
<td>RNA metabolism</td>
<td>89,90</td>
</tr>
<tr>
<td>MATR3</td>
<td>5q31.2</td>
<td>AD</td>
<td>WES</td>
<td>Established</td>
<td>ALS, FTD/ALS</td>
<td>RNA metabolism</td>
<td>95,222</td>
</tr>
<tr>
<td>hnRNPA1</td>
<td>12q13.13</td>
<td>AD</td>
<td>WES</td>
<td>To be confirmed</td>
<td>ALS, ALS/FTD, FTD</td>
<td>RNA metabolism</td>
<td>100</td>
</tr>
<tr>
<td>hnRNPA2B1</td>
<td>7p15.2</td>
<td>AD</td>
<td>WES</td>
<td>To be confirmed</td>
<td>ALS, ALS/FTD, FTD</td>
<td>RNA metabolism</td>
<td>100</td>
</tr>
<tr>
<td>ATXN2</td>
<td>12q24</td>
<td>NA</td>
<td>Candidate gene</td>
<td>Established</td>
<td>ALS</td>
<td>RNA translation, endocytosis</td>
<td>106</td>
</tr>
<tr>
<td>C9orf72</td>
<td>9q21.2</td>
<td>AD</td>
<td>GWAS, linkage</td>
<td>Established</td>
<td>ALS, ALS, ALS/FTD, FTD</td>
<td>DENN protein</td>
<td>108,109</td>
</tr>
<tr>
<td>NFH</td>
<td>22q12.2</td>
<td>AD</td>
<td>Candidate gene, association</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Axonal transport</td>
<td>159</td>
</tr>
<tr>
<td>DCTN1</td>
<td>2p13.1</td>
<td>AD</td>
<td>Linkage</td>
<td>Established</td>
<td>ALS, ALS/FTD</td>
<td>Axonal transport</td>
<td>162</td>
</tr>
<tr>
<td>PFN1</td>
<td>17p13</td>
<td>AD</td>
<td>WES</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Cytoskeleton</td>
<td>164</td>
</tr>
<tr>
<td>VAPB</td>
<td>20q13.33</td>
<td>AD</td>
<td>Linkage</td>
<td>Established</td>
<td>ALS</td>
<td>Vescle trafficking</td>
<td>172</td>
</tr>
<tr>
<td>Alsin</td>
<td>2q33.2</td>
<td>AR</td>
<td>Linkage</td>
<td>Established</td>
<td>Juvenile ALS</td>
<td>Vescle trafficking</td>
<td>173,226</td>
</tr>
<tr>
<td>TUBA4A</td>
<td>2q35</td>
<td>AD*</td>
<td>WES</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Cytoskeleton</td>
<td>175</td>
</tr>
<tr>
<td>CHCHD10</td>
<td>22q11.2</td>
<td>AD and AR</td>
<td>WES</td>
<td>Established</td>
<td>ALS, ALS/FTD, FTD</td>
<td>Oxidative phosphorylation</td>
<td>198,199</td>
</tr>
<tr>
<td>PPARG1A</td>
<td>4p15.2</td>
<td>NA</td>
<td>Candidate gene</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Mitochondrial biogenesis</td>
<td>204</td>
</tr>
<tr>
<td>Peripherin</td>
<td>12q12</td>
<td>AD*</td>
<td>Candidate gene</td>
<td>Uncertain</td>
<td>ALS</td>
<td>Cytoskeleton</td>
<td>224</td>
</tr>
<tr>
<td>FIG4</td>
<td>6q21</td>
<td>AD and AR</td>
<td>Candidate gene</td>
<td>Established</td>
<td>ALS</td>
<td>Vescle trafficking</td>
<td>227</td>
</tr>
<tr>
<td>SFG1I</td>
<td>15q21.1</td>
<td>AR</td>
<td>Linkage, candidate gene</td>
<td>To be confirmed</td>
<td>Juvenile ALS</td>
<td>Axonal maintenance, vesicle trafficking</td>
<td>228</td>
</tr>
<tr>
<td>SETX</td>
<td>9q34.13</td>
<td>AR</td>
<td>Linkage</td>
<td>Established</td>
<td>ALS</td>
<td>DNA/RNA metabolism</td>
<td>229</td>
</tr>
<tr>
<td>EWSR1</td>
<td>22q12.2</td>
<td>AD*</td>
<td>Candidate gene</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>RNA metabolism</td>
<td>230</td>
</tr>
<tr>
<td>TAF15</td>
<td>17q12</td>
<td>AD and AR</td>
<td>Candidate gene</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>RNA metabolism</td>
<td>231</td>
</tr>
<tr>
<td>DAO</td>
<td>13q33.2</td>
<td>AD</td>
<td>Linkage</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Oxidative deamination</td>
<td>232</td>
</tr>
<tr>
<td>ANG</td>
<td>14q11.1</td>
<td>AD*</td>
<td>Candidate gene</td>
<td>Established</td>
<td>ALS, ALS/FTD</td>
<td>Angiogenic factor</td>
<td>235</td>
</tr>
<tr>
<td>TBK1</td>
<td>12q14</td>
<td>AD</td>
<td>WES</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Autophagy</td>
<td>237</td>
</tr>
<tr>
<td><strong>Modifiers, susceptibility genes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATXN2</td>
<td>12q24</td>
<td>NA</td>
<td>Candidate gene</td>
<td>Established</td>
<td>ALS</td>
<td>RNA translation, endocytosis</td>
<td>106</td>
</tr>
<tr>
<td>EPHA4</td>
<td>2q36.1</td>
<td>NA</td>
<td>Candidate gene</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Axon guidance</td>
<td>168</td>
</tr>
<tr>
<td>UNC13A</td>
<td>19p13.11</td>
<td>NA</td>
<td>GWAS</td>
<td>Established</td>
<td>ALS, ALS/FTD</td>
<td>Neurotransmitter release</td>
<td>184,185,220</td>
</tr>
<tr>
<td>ELP3</td>
<td>8p21.1</td>
<td>NA</td>
<td>GWAS</td>
<td>Established</td>
<td>ALS, ALS/FTD</td>
<td>RNA metabolism</td>
<td>188,192</td>
</tr>
<tr>
<td>SS18L1</td>
<td>20q13.33</td>
<td>AD</td>
<td>WES</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Chromatin regulation</td>
<td>203</td>
</tr>
<tr>
<td>APOE</td>
<td>19q13.32</td>
<td>NA</td>
<td>Candidate gene</td>
<td>To be confirmed</td>
<td>ALS, FTD</td>
<td>Lipid homeostasis</td>
<td>207,219</td>
</tr>
<tr>
<td>KIFAP3</td>
<td>1q24.2</td>
<td>NA</td>
<td>GWAS</td>
<td>Uncertain</td>
<td>ALS</td>
<td>Axonal transport</td>
<td>207</td>
</tr>
<tr>
<td>VEGF</td>
<td>6p21.1</td>
<td>NA</td>
<td>Candidate gene</td>
<td>Uncertain</td>
<td>ALS</td>
<td>Angiogenic factor</td>
<td>209</td>
</tr>
<tr>
<td>SMN1</td>
<td>5q13.2</td>
<td>NA</td>
<td>Candidate gene</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Survival factor</td>
<td>221</td>
</tr>
<tr>
<td>ZNFS12B</td>
<td>20q13.33</td>
<td>NA</td>
<td>GWAS</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Positive regulator of TGF-β signaling</td>
<td>223</td>
</tr>
<tr>
<td>SPAST</td>
<td>2q22.3</td>
<td>AD</td>
<td>Candidate gene</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Cytoskeleton</td>
<td>225</td>
</tr>
</tbody>
</table>

(Continued)
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Genes</th>
<th>Locus</th>
<th>Inheritance</th>
<th>Discovery method</th>
<th>Suggested role in ALS*</th>
<th>Involvement in the ALS-FTD spectrum</th>
<th>Putative protein function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>PON1, 2, 3</td>
<td>7q21.3</td>
<td>NA</td>
<td>Candidate gene</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Detoxifying enzyme</td>
<td>234</td>
</tr>
<tr>
<td>HFE</td>
<td>6p22.2</td>
<td>NA</td>
<td>Candidate gene</td>
<td>To be confirmed</td>
<td>ALS</td>
<td>Iron metabolism</td>
<td>235</td>
</tr>
<tr>
<td>TMEM106B</td>
<td>7p21.3</td>
<td>NA</td>
<td>GWAS</td>
<td>Established</td>
<td>FTD</td>
<td>Endolysosomal pathway</td>
<td>236</td>
</tr>
</tbody>
</table>

Notes: "The suggested role of the different genes in ALS pathogenesis is an appreciation based upon literature and following criteria: “Established”: generally accepted as causal, modifying, or susceptibility genes in independent studies; “To be confirmed”: firm scientific basis but further validation in other populations is mandatory; “Uncertain”: further study is needed. Further evidence is required.

Abbreviations: ALS, amyotrophic lateral sclerosis; FTD, frontotemporal dementia; AD, autosomal dominant; AR, autosomal recessive; XL, X-linked inheritance; GwAS, genome-wide association study; WES, whole-exome sequencing; TGF-β, transforming growth factor-β; NA, not applicable.

superoxide radicals and protects the cell against reactive oxygen species. So far, over 160 mutations scattered throughout all five exons have been identified (ALSod Consortium; http://alsod.iop.kcl.ac.uk/), predominantly consisting of missense mutations; nonsense mutations or gene deletions/insertions appear to be rare.7 The majority of mutations are autosomal dominantly inherited, with the exception of the recessive D90A mutation in the Scandinavian population.34

![Schematic overview of processes that contribute to ALS pathogenesis.](https://www.dovepress.com/)

**Figure 1** Schematic overview of processes that contribute to ALS pathogenesis.

**Notes:** (A) Collapse of the proteostasis machinery: impairment of normal proteasomal or autophagic degradation is caused by mutations in, among others, ubiquilin 2 (UBQLN2), valosin-containing protein (VCP), sequestosome 1 (SQSTM1), charged multivesicular body protein 2b (CHMP2B), optineurin (OPTN), and superoxide dismutase 1 (SOD1). By interfering with the proteostatic machinery, protein aggregates fuel a self-reinforcing vicious circle of misfolded protein buildup and subsequent disturbance of various cellular processes; (B) RNA-mediated motor neuron degeneration: disturbance of normal RNA processing, which results in abnormally assembled proteins and toxic RNA species, is caused by mutations in, among others, fused in sarcoma (FUS), TAR DNA-binding protein (TARDBP), C9orf72, matrin3 (MATR3), and heterogeneous nuclear ribonucleoprotein complex proteins (hnRNPs); (C) axonal dysfunction: mutations in, among others, tubulin alpha 4A (TUBA4A), dynactin, peripherin, neurofilament heavy chain (NFH), and profilin 1 (PFN1) affect the axonal architecture and function.

**Abbreviations:** ALS, amyotrophic lateral sclerosis; Ub, ubiquitin.
For some mutations, genotype–phenotype correlations have been established. For example, carriers of the A4V mutation clinically mainly show lower motor neuron involvement and suffer from an aggressive ALS form with rapid disease progression and death on average occurring at 1.4 years after symptom onset. In contrast, D90A homozygous patients have a rather mild phenotype with mean disease duration of around 14 years. It remains to be seen whether all reported SOD1 variants are truly pathogenic.

The identification of SOD1 mutations has meant a momentous turnaround for ALS research. Rodent models overexpressing mutant SOD1 characterized by a fatal motor neuron degeneration reminiscent of ALS in humans were developed. In particular, the SOD1[G93A] mouse model has been widely used in order to get insight into the pathogenic mechanism of mutant SOD1. Several hypotheses regarding the mechanisms by which SOD1 aggregates contribute to the manifestation of motor neuron death have been proposed. These include perturbation of mitochondrial function, glutamate excitotoxicity, disturbance of axonal transport, SOD1 toxicity in nonneuronal cells surrounding the motor neurons, and impairment of protein homeostasis.

One of the contributing mechanisms of mutant SOD1 may relate to protein homeostasis. Evidence suggests that the cell’s chaperone system is unable to refold misfolded mutant SOD1, resulting in abundant presence of toxic (likely oligomeric) SOD1 species. Consequently, the ubiquitin–proteasome system (UPS) and autophagy pathway appear to be unable to dispose these. In spinal cord motor neurons of ALS patients and rodent models, increased autophagosome expression was found, whereas proteasome activity was reduced. Moreover, the misfolded proteins impair the protein degradation machinery in the cell, which leads to a vicious circle of misfolded protein buildup through which a variety of important cellular processes are perturbed, resulting in axonal retraction and neuronal cell death.

Still, even after 20 years of mutant SOD1 research, an unequivocal explanation for its toxicity has yet to be provided; a combination of different mechanisms is likely. Furthermore, most of data available in the literature relate to what happens in the mutant SOD1 mouse. Evidence for what happens in human ALS is less astounding.

Of notice, posttranslational modifications cause wild-type SOD1 to misfold and adopt an aggregation-prone and toxic conformation comparable to mutant SOD1. Using a conformation-specific antibody, aberrantly folded wild-type SOD1 has been described in the motor neurons in a subset of SALS cases (thus, without SOD1 mutations). These observations implicate that aberrantly modified forms of wild-type SOD1 may be at play in SALS pathogenesis.

### Disturbance of protein degradation in ALS: a common theme

Strong evidence for a disturbance of protein degradation as a cellular disease mechanism in ALS comes from the exciting finding that ALS can be caused by mutations in proteins, which are directly or indirectly involved in regulated protein breakdown (Figure 2A).

#### Ubiquilin 2 (UBQLN2)

In 2011, missense mutations in the UBQLN2 gene were found to underlie X-linked ALS and ALS/FTD. They appear to be rare. UBQLN2 is a member of the ubiquitin-like (UBL) protein family and consists of a C-terminal ubiquitin-associated (UBA) domain together with an N-terminal UBL domain. Through these domains, UBQLN2 binds poly-ubiquitin chains and associates with the proteasome and autophagosome. It thus mediates the delivery of ubiquitin-tagged proteins for degradation. Interestingly, the most prevalent UBQLN2 mutations affect the proline residue in a region containing 12 PXX tandem repeats (P497H, P497S, P506T, P509S, and P525S), but mutations outside this region have also been described. Of notice, UBQLN2 can also be found in skein-like inclusions in the spinal cord motor neurons and hippocampus of ALS and ALS/FTD patients not carrying mutations in this protein, suggesting that UBQLN2 may have a role in ALS that is not only attributable to mutations.

#### Sequestosome 1 (SQSTM1)

SQSTM1, also known as p62, shares structural similarity with UBQLN2 and is found in inclusions in a variety of neurodegenerative disorders. The protein acts as a cargo receptor for the degradation of ubiquitinated proteins and regulates the activation of nuclear factor kappa-B (NF-κB) signaling. SQSTM1 mutations have been found in SALS, FALS, FTD, and FTD/ALS patients. Further screening and functional studies are needed to understand the pathogenic relevance of these SQSTM1 alterations for ALS.

#### Valosin-containing protein (VCP)

Exome sequencing revealed mutations in the VCP gene in ~1%–2% of FALS and <1% of SALS patients. Such mutations had already been found in a syndrome characterized by inclusion body myopathy (IBM) with Paget’s disease of the bone (PDB) and FTD (IBMPFD). VCP (also known as p97) is a multifunctional AAA++-ATPase that orchestrates...
Advances in Genomics and Genetics 2015:5

Figure 2 Schematic overview of selected ALS genes according to their (neuronal) function.

Notes: (A) Protein degradation: delivery of misfolded and ubiquitinated (Ub) proteins to the proteasome and their subsequent degradation is facilitated by binding of sequestosome 1 (SQSTM1), ubiquitin 2 (UBQLN2), or valosin-containing protein (VCP). In addition, VCP is essential for the maturation of autophagosomes and facilitates the autophagic clearance of stress granules. Optineurin (OPTN) and TANK-binding kinase 1 (TBK1) are involved in the autophagic clearance of protein aggregates. Charged multivesicular body protein 2b (CHMP2B) is required for the sorting of integral membrane proteins in multivesicular bodies and their subsequent autophagic degradation. Superoxide dismutase 1 (SOD1) aggregates impair the protein degradation machinery in the cell, leading to a vicious circle of misfolded protein buildup through which a variety of important cellular processes are perturbed. (B) DNA/RNA metabolism: disturbance of normal RNA processing and metabolism results in abnormally assembled proteins and toxic RNA species. To this extent, the C9orf72 RNA repeat expansion may exert its toxic properties via the sequestration of multiple RNA-containing proteins in RNA foci and their subsequent depletion. Another possible disease mechanism is the contribution of repeat-associated non-ATG (RAN) translation products, more specifically dipeptide proteins (DPRs). Ataxin-2 (ATXN2) intermediate-length polyQ expansions and aggregates may lead to depletion of the protein and therefore may inhibit the reported functions of ATXN2 in RNA metabolism. Other regulators of RNA homeostasis that are mutated in ALS are senatin (SETX) and Ewing sarcoma breakpoint region 1 (EWSR1). Elongator acetyltransferase complex subunit 3 (Elongator complex subunit 3, Elongator subunit 3) is involved in RNA processing through mRNA elongation, histone acetylation, and modification of transfer RNA (tRNA) wobble nucleosides. Matrin 3 (MATR3) is predominantly associated with the nuclear matrix and is involved in RNA metabolism. SS18-like protein 1 (SS18L1) is a chromatin regulatory protein that is a component of the neuronal chromatin remodeling complex (nBAF). TAF15 has similar in vitro and in vivo properties to TDP-43 and is a component of the neuronal chromatin remodeling complex (nBAF). (C) Cytoskeletal function: neurofilaments (assembled from light chain 3; Ac, acetyl group; DPR, dipeptide ribosomes) and their subsequent depletion. Another possible disease mechanism is the contribution of repeat-associated non-ATG (RAN) translation products, more specifically dipeptide proteins (DPRs). Ataxin-2 (ATXN2) intermediate-length polyQ expansions and aggregates may lead to depletion of the protein and therefore may inhibit the reported functions of ATXN2 in RNA metabolism. Other regulators of RNA homeostasis that are mutated in ALS are senatin (SETX) and Ewing sarcoma breakpoint region 1 (EWSR1). Elongator acetyltransferase complex subunit 3 (Elongator complex subunit 3, Elongator subunit 3) is involved in RNA processing through mRNA elongation, histone acetylation, and modification of transfer RNA (tRNA) wobble nucleosides. Matrin 3 (MATR3) is predominantly associated with the nuclear matrix and is involved in RNA metabolism. SS18-like protein 1 (SS18L1) is a chromatin regulatory protein that is a component of the neuronal chromatin remodeling complex (nBAF). TAF15 has similar in vitro and in vivo properties to TDP-43 and is a component of the neuronal chromatin remodeling complex (nBAF). (C) Cytoskeletal function: neurofilaments (assembled from light chain 3; Ac, acetyl group; DPR, dipeptide ribosomes) and their subsequent depletion. Another possible disease mechanism is the contribution of repeat-associated non-ATG (RAN) translation products, more specifically dipeptide proteins (DPRs). Ataxin-2 (ATXN2) intermediate-length polyQ expansions and aggregates may lead to depletion of the protein and therefore may inhibit the reported functions of ATXN2 in RNA metabolism. Other regulators of RNA homeostasis that are mutated in ALS are senatin (SETX) and Ewing sarcoma breakpoint region 1 (EWSR1). Elongator acetyltransferase complex subunit 3 (Elongator complex subunit 3, Elongator subunit 3) is involved in RNA processing through mRNA elongation, histone acetylation, and modification of transfer RNA (tRNA) wobble nucleosides. Matrin 3 (MATR3) is predominantly associated with the nuclear matrix and is involved in RNA metabolism. SS18-like protein 1 (SS18L1) is a chromatin regulatory protein that is a component of the neuronal chromatin remodeling complex (nBAF). TAF15 has similar in vitro and in vivo properties to TDP-43 and is a component of the neuronal chromatin remodeling complex (nBAF). (C) Cytoskeletal function: neurofilaments (assembled from light chain 3; Ac, acetyl group; DPR, dipeptide ribosomes) and their subsequent depletion. Another possible disease mechanism is the contribution of repeat-associated non-ATG (RAN) translation products, more specifically dipeptide proteins (DPRs). Ataxin-2 (ATXN2) intermediate-length polyQ expansions and aggregates may lead to depletion of the protein and therefore may inhibit the reported functions of ATXN2 in RNA metabolism. Other regulators of RNA homeostasis that are mutated in ALS are senatin (SETX) and Ewing sarcoma breakpoint region 1 (EWSR1). Elongator acetyltransferase complex subunit 3 (Elongator complex subunit 3, Elongator subunit 3) is involved in RNA processing through mRNA elongation, histone acetylation, and modification of transfer RNA (tRNA) wobble nucleosides. Matrin 3 (MATR3) is predominantly associated with the nuclear matrix and is involved in RNA metabolism. SS18-like protein 1 (SS18L1) is a chromatin regulatory protein that is a component of the neuronal chromatin remodeling complex (nBAF). TAF15 has similar in vitro and in vivo properties to TDP-43 and is a component of the neuronal chromatin remodeling complex (nBAF). (C) Cytoskeletal function: neurofilaments (assembled from light chain 3; Ac, acetyl group; DPR, dipeptide ribosomes) and their subsequent depletion. Another possible disease mechanism is the contribution of repeat-associated non-ATG (RAN) translation products, more specifically dipeptide proteins (DPRs). Ataxin-2 (ATXN2) intermediate-length polyQ expansions and aggregates may lead to depletion of the protein and therefore may inhibit the reported functions of ATXN2 in RNA metabolism. Other regulators of RNA homeostasis that are mutated in ALS are senatin (SETX) and Ewing sarcoma breakpoint region 1 (EWSR1). Elongator acetyltransferase complex subunit 3 (Elongator complex subunit 3, Elongator subunit 3) is involved in RNA processing through mRNA elongation, histone acetylation, and modification of transfer RNA (tRNA) wobble nucleosides. Matrin 3 (MATR3) is predominantly associated with the nuclear matrix and is involved in RNA metabolism. SS18-like protein 1 (SS18L1) is a chromatin regulatory protein that is a component of the neuronal chromatin remodeling complex (nBAF). TAF15 has similar in vitro and in vivo properties to TDP-43 and is a component of the neuronal chromatin remodeling complex (nBAF). (C) Cytoskeletal function: neurofilaments (assembled from light chain 3; Ac, acetyl group; DPR, dipeptide ribosomes) and their subsequent depletion. Another possible disease mechanism is the contribution of repeat-associated non-ATG (RAN) translation products, more specifically dipeptide proteins (DPRs). Ataxin-2 (ATXN2) intermediate-length polyQ expansions and aggregates may lead to depletion of the protein and therefore may inhibit the reported functions of ATXN2 in RNA metabolism. Other regulators of RNA homeostasis that are mutated in ALS are senatin (SETX) and Ewing sarcoma breakpoint region 1 (EWSR1). Elongator acetyltransferase complex subunit 3 (Elongator complex subunit 3, Elongator subunit 3) is involved in RNA processing through mRNA elongation, histone acetylation, and modification of transfer RNA (tRNA) wobble nucleosides. Matrin 3 (MATR3) is predominantly associated with the nuclear matrix and is involved in RNA metabolism. SS18-like protein 1 (SS18L1) is a chromatin regulatory protein that is a component of the neuronal chromatin remodeling complex (nBAF). TAF15 has similar in vitro and in vivo properties to TDP-43 and is a component of the neuronal chromatin remodeling complex (nBAF). (C) Cytoskeletal function: neurofilaments (assembled from light chain 3; Ac, acetyl group; DPR, dipeptide ribosomes) and their subsequent depletion. Another possible disease mechanism is the contribution of repeat-associated non-ATG (RAN) translation products, more specifically dipeptide proteins (DPRs). Ataxin-2 (ATXN2) intermediate-length polyQ expansions and aggregates may lead to depletion of the protein and therefore may inhibit the reported functions of ATXN2 in RNA metabolism. Other regulators of RNA homeostasis that are mutated in ALS are senatin (SETX) and Ewing sarcoma breakpoint region 1 (EWSR1). Elongator acetyltransferase complex subunit 3 (Elongator complex subunit 3, Elongator subunit 3) is involved in RNA processing through mRNA elongation, histone acetylation, and modification of transfer RNA (tRNA) wobble nucleosides. Matrin 3 (MATR3) is predominantly associated with the nuclear matrix and is involved in RNA metabolism. SS18-like protein 1 (SS18L1) is a chromatin regulatory protein that is a component of the neuronal chromatin remodeling complex (nBAF). TAF15 has similar in vitro and in vivo properties to TDP-43 and is a component of the neuronal chromatin remodeling complex (nBAF). (C) Cytoskeletal function: neurofilaments (assembled from light chain 3; Ac, acetyl group; DPR, dipeptide ribosomes) and their subsequent depletion. Another possible disease mechanism is the contribution of repeat-associated non-ATG (RAN) translation products, more specifically dipeptide proteins (DPRs). Ataxin-2 (ATXN2) intermediate-length polyQ expansions and aggregates may lead to depletion of the protein and therefore may inhibit the reported functions of ATXN2 in RNA metabolism.

Abbreviations: ALS, amyotrophic lateral sclerosis; ER, endoplasmic reticulum; LC3, microtubule-associated protein 1A/1B-light chain 3; Ac, acetyl group; DPR, dipeptide protein; RRM, RNA recognition motif.
a number of processes, including autophagy and proteasomal degradation. More precisely, VCP directs ubiquitintagged proteins to the proteasome and is essential for the maturation of autophagosomes. Interestingly, nearly all mutations reside in the N-terminal domain, which is crucial to carry out the proper function of VCP. Moreover, in mammalian cells, the depletion of pathogenic mutations in VCP reduced the autophagic clearance of stress granules, which are presumed to create a proaggregation environment. Therefore, it is plausible that VCP mutations disrupt VCP’s protein and stress granule removal activity in ALS, giving rise to the accumulation of ubiquitinated protein deposits within the cell.

**Optineurin (OPTN)**

OPTN mutations are associated with a very diverse phenotype. Polymorphisms have been described in Paget’s disease; missense and insertion mutations were found in primary open-angle glaucoma (POAG); and OPTN mutations have now been reported as a (rare) cause of ALS. OPTN is a protein with many functions; it is involved in membrane and vesicle trafficking, NF-κB regulation, and autophagic clearance of protein aggregates. Homozygous deletions, nonsense mutations, and heterozygous missense mutations have been found in both FALS and SALS patients. OPTN depletion enhances protein aggregation in Hela cells, and OPTN inclusions are found in the spinal cord of SALS patients. The exact mechanism through which OPTN mutations cause ALS still remains to be established.

**Charged multivesicular body protein 2b (CHMP2B)**

Although rarely so, mutations in CHMP2B originally found to be associated with FTD may give rise to an ALS phenotype. CHMP2B is required for the sorting of integral membrane proteins in multivesicular bodies and their subsequent autophagic degradation. The finding of ALS-causing mutations in these proteins suggests that disturbances of protein degradation play a mechanistic role in ALS, as mentioned earlier. However, it should be noted that almost all mutations are missense mutations giving rise to proteins that, at most, have lost some of their function. This means that in the cell (apart from the UBQLN2 mutations), at least 50% of the normal function is preserved. Therefore, it remains to be convincingly shown that these mutations actually cause failure of the protein breakdown sufficient to induce cell dysfunction and death. It is fairly well possible that they are pathogenic not through such a loss-of-function mechanism, but rather through a gain-of-function mechanism such as described for SOD1 mutants. Obviously, both mechanisms may be at play.

**ALS: dysregulation of RNA householding**

The identification of disease-linked mutations in genes involved in RNA processing marked the beginning of a paradigm shift and challenged the proteocentric thinking of the ALS field. May aberrant RNA metabolism contribute to ALS pathogenesis (Figure 2B)?

**TAR DNA-binding protein (TARDBP)**

The TARBP (TDP-43) is a major constituent of ubiquitin-included inclusions in a variety of neurodegenerative disorders, including FTD and ALS. This observation prompted the mutational screening of the TARDBP gene (which encodes the TDP-43 protein) in a cohort of FALS and SALS cases. To this day, up to 47 missense mutations and one truncated variant have been identified. Nearly all of the missense mutations reside in exon 6, which encodes the C-terminal glycine-rich part. TARDBP mutations have been reported in 4% of FALS and 1% of SALS cases. TDP-43 is a multifunctional nuclear protein that is, among other things, involved in multiple levels of RNA processing including transcription, splicing, transport, and translation. Besides the C-terminal glycine-rich domain, TDP-43 also contains two RNA recognition motifs (RRMs). RM1 is essential for binding to single-stranded RNA with a minimum of five GU repeats. GU-rich sequences are predominantly present in pre-mRNAs encoded by more than 6,000 genes, which are consequently targeted by TDP-43. Therefore, it is not surprising that depletion of TDP-43 with the aid of antisense oligonucleotides affected the expression of ~600 mRNAs and altered 965 splicing events in the mouse brain. Of note, long intronic sequences containing GU repeats are typically found in neuronally expressed genes, and thus many of these can be found on the list of genes affected by TDP-43 downregulation.

Although TDP-43 is predominantly present in the nucleus, nuclear-cytoplasmic shuttling has been shown, a process controlled by its bipartite nuclear localization signal (NLS) and nuclear export signal (NES). Neurons and glial cells of the ALS nervous system display a redistribution of TDP-43 from the nucleus to the cytoplasm. In the cytoplasmic compartment, TDP-43 associates with stress granules that appear in times of cellular stress (Figure 2B). In these granules, mRNAs are sorted during a triage process that determines...
mRNA fate: translation, sequestration in the stress granule, or degradation. Stress granules in ALS presumably create a proaggregation environment and hereby contribute to the observed inclusion build up and TDP-43 depletion from the nucleus.

Based upon this, a multistep model has been suggested to explain TDP-43 toxicity. First, ALS-associated mutations augment cytoplasmic TDP-43 accumulation (loss-of-function in the nucleus) and enhance its aggregation propensity (gain-of-function). Second, stress granule formation further reinforces TDP-43 accumulation and aggregation. Their toxicity may further be enhanced by acting as a sink for other proteins and RNAs (loss-of-function). It is likely, based on what has been found in cellular and animal models, that both TDP-43 deficiency and overexpression are detrimental for the living organism. However, in at least one mouse model, TDP-43 overexpression induced a phenotype without any evidence for aggregate formation or TDP-43 mislocalization. The understanding of the interplay between TDP-43 aggregation, stress granules, and TDP-43 pathology needs further clarification.

Fused in sarcoma (FUS)
The interest in RNA processing as a mechanistic factor in motor neuron degeneration grew rapidly with the discovery of FUS mutations. FUS shares structural homology with TDP-43 and is also involved in RNA metabolism. The protein binds over 5,500 genes through a GUGGU-binding motif and FUS depletion alters the splicing of more than 950 mRNAs. Remarkably, most of them are distinct from the mRNAs targeted by TDP-43. Similarly to what is seen in TDP-43, nearly all mutations are clustered in the C-terminal region encoded by exons 14 and 15. FUS variants account for ~4% of FALS and ~1% of SALS. Cognitive impairment is rare in ALS patients with FUS mutations and the overall survival is significantly shorter compared with other ALS forms. Of note, a possible association between FUS mutations (c. 1475delG and c. 1542G > T) and mental retardation was recently discovered.

Under physiological conditions, FUS is predominantly present in the cell nucleus. Similar to what is seen for TDP-43, cytoplasmic mislocalization and FUS-positive inclusions can be found in ALS patients with FUS mutations. It is thought that this is explained by mutations in the NLS, which compromise transportin-mediated nuclear import. Interference with this transport results in cytoplasmic redistribution and subsequent recruitment of FUS protein into stress granules (Figure 2B). As stress granules are known to be stress-sensitive dynamic entities, it seems possible that they form insoluble aggregates under stress conditions. A similar multistep model as for TDP-43 toxicity has been proposed for FUSopathies. Again, it remains to be seen whether a loss-of-function or gain-of-function mechanism or both explain what happens in patients.

Matrin3 (MATR3)
In early 2014, Johnson et al identified missense mutations in MATR3 that segregated with disease in several families with multiple members affected by ALS. Interestingly, a MATR3 variant (Ser85Cys) had been found to underlie a distal, asymmetrical myopathy with vocal cord weakness in two large families of different origin but it turns out that these patients actually suffered from a slowly progressive motor neuron disorder better classified as ALS.

MATR3 is a highly conserved protein that is predominantly associated with the nuclear matrix. It has structural similarity with other ALS-linked genes, ie, TDP-43 and FUS. With its bipartite NLS, two zinc finger domains, and two RRM s, MATR3 is involved in RNA metabolism. It binds and stabilizes several mRNA species including that of TDP-43. Interestingly, one MATR3 variant (Ser85Cys) displayed a higher affinity for TDP-43 while its interaction with other partners remained unaltered. Cytoplasmic mislocalization of mutant MATR3 is found in motor neurons of some ALS patients, even without these patients carrying MATR3 mutations. It is tempting to speculate that MATR3 variants cause dysfunction of TDP-43. Despite being a rare cause of ALS, the discovery of MATR3 mutations underscores once again the importance of RNA metabolism in ALS.

Heterogeneous nuclear ribonucleoprotein complex proteins (hnRNPs): infectious proteins?
Recently, mutations in the prion-like domain of hnRNP2/B1 and hnRNP1 have been reported in patients suffering from ALS and IBMPFD. Notably, these mutations enhance hnRNP2/B1 and hnRNP1 incorporation into stress granules and exacerbate their intrinsic tendency to assemble into self-seeding fibrils. Multiple lines of evidence link hnRNPs to neurodegeneration. They directly interact with TDP-43 to regulate RNA metabolism and two of them (hnRNP2/B1 and hnRNP1) are suppressors of VCP-induced degeneration in a Drosophila model for multisystem proteinopathy. hnRNP2/B1 and hnRNP1 mutations are rare in ALS. Finding more families in which they segregate with the disease would be reassuring, but
their identification has contributed to the generation of a hypothesis for a prion-like mechanism contributing to ALS.\textsuperscript{26} Pathogenic mutations strengthen a steric zipper motif in the prion-like domain of hnRNP proteins, which accelerates the formation of self-seeding fibrils that induce the polymerization of wild-type hnRNPs.\textsuperscript{106} It is thought that such motifs may induce aggregation by acting as a template, forcing the conversion of natively folded proteins into an abnormal form. The resulting aggregates may be released and taken up by neighboring cells, thus explaining the spatially progressive nature of ALS.\textsuperscript{44,103} Interestingly, not only hnRNP proteins, but also SOD1, TDP-43, and FUS have been shown to contain such a prion-like domain.\textsuperscript{28,104} It should be noted that these findings strongly link RNA householding dysregulation and dysproteostasis as two pathogenic pathways for motor neuron degeneration.

### C9orf72 mutations in ALS: secrets unlocked?

Since the discovery of repeat expansions as a mechanism of disease in the 1990s, at least 24 neurological disorders characterized by such mutations have been described.\textsuperscript{105} Repeat expansions in the Ataxin-2 (ATXN2) gene\textsuperscript{106} and the non-imprinted Prader–Willi/Angelman syndrome region protein 1 (NIPA1) gene\textsuperscript{107} had previously already been implicated in ALS and/or FTD. But recently, major progress in the field of ALS was made by the discovery of a noncoding hexanucleotide repeat expansion (G\textsubscript{4}C\textsubscript{2}) in the 5’ region of C9orf72 as a very prevalent cause of ALS and FTD.\textsuperscript{108,109} While in the normal population the number of repeats never exceeds 30, their number in C9orf72 mutation carriers increases to up to thousand repeats.\textsuperscript{108,109}

The exact function of the cytoplasmic C9orf72 protein remains to be elucidated at present. The structural similarity to DENN-like proteins suggested that the protein may belong to this family of GDP–GTP exchange factors for Rab-GTPases, which are involved in the regulation of membrane trafficking.\textsuperscript{110} At least three different C9orf72 transcripts exist and are expressed in most tissues including brain.\textsuperscript{108}

An impressive percentage of ALS and FTD can currently be explained by C9orf72 mutations: nearly 40% of FALS and almost 10% of SALS; more than 25% of familial FTD; and approximately 5% of sporadic FTD.\textsuperscript{111} The G\textsubscript{4}C\textsubscript{2} expansion is associated with a variable phenotype as addressed in recent reviews.\textsuperscript{2,24,111} C9orf72-ALS has a high incidence of bulbar onset and cognitive dysfunction is frequent with C9orf72 mutations.\textsuperscript{2,24,111} Almost a third of C9orf72 patients have both ALS and FTD.\textsuperscript{111} Additional phenotypes include Parkinsonism, ataxia, and psychosis.\textsuperscript{112,113} In addition, a large variation in disease onset (27–83 years) and duration (3–264 months) is evident among C9orf72 subjects.\textsuperscript{111}

Interestingly, some C9orf72 patients harbor variations in other ALS- and/or FTD-associated genes,\textsuperscript{111,114} a finding of possible pathogenic significance.\textsuperscript{111} It is thought, but far from proven, that the C9orf72 expansion mutation establishes a susceptibility for neurodegeneration and that additional factors (such as these mutations or other modifiers) shape the phenotype, acting as disease modifiers.\textsuperscript{111} Of note, in that regard, the motor component appears to be modified by genes different from those that modify the cognitive component.\textsuperscript{111}

Nuclear RNA foci containing both sense\textsuperscript{108,115–117} and antisense\textsuperscript{116,118,119} repeats have been found in neurons of the frontal cortex, and the spinal cord of patients with C9orf72 mutations.\textsuperscript{108} Three plausible mechanisms have been proposed to explain the pathogenic effect of the C9orf72 expansions and they may exist simultaneously.

First, the expansion may interfere with C9orf72 expression. Several, but certainly not all studies have found a reduction of mRNA expression in patients with the C9orf72 expansion, possibly resulting in haploinsufficiency (loss-of-function).\textsuperscript{120,121} Motor axonal degeneration was indeed seen upon knockdown of C9orf72 in zebrafish.\textsuperscript{122} In addition, hypermethylation of the C9orf72 promoter in C9orf72 mutation carriers was found.\textsuperscript{123} This methylation-driven gene silencing also occurs in other repeat disorders, including Friedreich’s ataxia,\textsuperscript{124} fragile X mental retardation,\textsuperscript{125} and myotonic dystrophy (DM).\textsuperscript{126–128} However, promoter hypermethylation was associated with reduced accumulation of RNA foci and dipeptide repeat (DPR) protein aggregates in C9orf72 patient brains\textsuperscript{129} and a shorter disease duration,\textsuperscript{130} suggesting that hypermethylation actually represents a protective response.\textsuperscript{129} Furthermore, homozygosity is not associated with more severe disease.\textsuperscript{130,131} These findings, together with the fact that C9orf72 coding mutations have not been described yet, make a loss-of-function mechanism less likely to have a major contribution to the molecular pathogenesis of C9orf72 expansion mutations.\textsuperscript{132}

In contrast, a body of evidence has emerged that favors a toxic RNA gain-of-function mechanism. The C9orf72 RNA repeat expansion may exert its toxic properties via the sequestration of multiple RRM-containing proteins and their subsequent depletion.\textsuperscript{117,133,134} Proteins involved in splicing, mRNA nuclear export, and/or translation were significantly enriched as binding partners of the G\textsubscript{4}C\textsubscript{2} expansion. hnRNPA3, hnRNPA2/B1, SFQP, ILF3, NONO, hnRNP L, IL2BP1, ILF-2, FUS,\textsuperscript{133} and Pur3\textsuperscript{135} displayed
strong and selective binding to the \(G_C\) repeat. Furthermore, immunohistochemical colocalization of RNA foci with SRSF2, hnRNPH1/F, and ALYREF in neuronal cells was observed.\(^{134}\) The avid and dynamic binding of these proteins to the expanded \(G_C\) sequence is presumably facilitated by the tendency of the repeat to form G-quadruplexes.\(^{136}\) Such deleterious gain-of-function of the \(G_C\) repeat expansion is similar to what is thought to occur in DM. This neuro muscular disorder is caused by a CTG repeat in the 3'UTR region of the myotonic dystrophy protein kinase (DMPK) gene (type one) or by a CCTG repeat in intron 1 of the zinc finger protein 9 (ZNF9) gene (type two).\(^{137-139}\) These repeats act as a sink for RNA binding proteins, resulting in the dysregulation of the (developmentally regulated) splicing of a variety of mRNAs such as the one of tau, the CIC-1 chloride channel, and troponin, explaining the clinical findings in DM patients.\(^{140-142}\) Sequestration of muscleblind-like 1 (MBNL1) appears to be crucial, as MBNL1-deficient mice display a phenotype reminiscent to that of DM.\(^{143,144}\)

Major attention has recently been devoted to yet another possible disease mechanism, the contribution of repeat-associated non-ATG (RAN) translation products (Figure 2B).\(^{145}\) This ATG-independent translation from both sense and antisense C9orf72 repeat transcripts gives rise to five different DPR species: poly-(Gly-Ala), poly-(Gly-Arg), poly-(Gly-Pro), poly-(Pro-Arg), and poly-(Pro-Ala).\(^{146,147}\) These DPRs are found in (TPD-43 negative, p62/ ubiquitin positive inclusions) the central nervous system of C9orf72 mutation carriers.\(^{147,148}\) When overexpressed in cells, yeast, or Drosophila, some of these DPRs evoke toxicity.\(^{149,150}\) Most but not all reports suggest poly-(Gly-Arg) and poly-(Pro-Arg) to be toxic, possibly through inducing nucleolar stress.\(^{148,151}\) It looks as if the \(G_C\) repeat toxicity in Drosophila eye is mediated by the generation of DPR species,\(^{150}\) but their pathological significance for the human condition still remains to be demonstrated. Still, even if their pathogenic contribution is limited, they may turn out to be very useful biomarkers.

**Cytoskeletal defects in ALS: an emerging theme**

With their axons extending more than one meter, motor neurons are highly dependent on axonal transport to shuttle organelles and vesicles between soma and synapses for their proper function and survival.\(^{152}\) Axonal swelling with cytoskeletal disarrangement is a hallmark of ALS pathology.\(^{153}\) Not surprisingly, a number of factors involved in axonal architecture and function have been documented to play a role in ALS (Figure 2C).\(^{154}\)

Neurofilaments are essential for normal nerve cell function and are assembled from light (NFL), medium (NFM), and heavy (NFH) subunits. Decreased expression levels of NFL (encoded by the NEFL gene) were previously demonstrated in ALS patients.\(^{155}\) Interestingly, TDP-43 and FUS bind to NEFL mRNA,\(^{75,91,156}\) which is sequestered in stress granules in ALS motor neurons.\(^{157}\) Furthermore, NFH mutations were found in ALS patients.\(^{158,159}\) Overexpression of the intermediate filament peripherin, which is a component of inclusion bodies associated with degenerating motor neurons, induces motor neuron abnormalities in mice,\(^{160}\) and a frameshift deletion in the peripherin gene has been reported in ALS.\(^{161}\) The significance of both types of mutations requires more study.

**Dynactin1 (DCTN1) mutations** have been identified in a lower motor neuron disease consisting of vocal fold paralysis\(^{162}\) and were later reported in FALS and SALS patients.\(^{163}\) DCTN1 stabilizes the binding of cargos to the motor protein dynein and DCTN1 variants may therefore contribute to the axonal transport deficits seen in ALS.

Notwithstanding, it is the recent discovery of profilin1 (PFN1) mutations as a cause for ALS that strongly points out the possible involvement of the cytoskeleton in ALS pathogenesis.\(^{164,165}\) PFN1 is an important regulator of actin dynamics.\(^{166}\) Mutant PFN1 may contribute to ALS by its reduced actin binding, thus diminished actin polymerization, and subsequent cytoskeletal disturbance.\(^{164}\) In addition, PFN1 mutations may link the cytoskeletal defects and RNA aggregation seen in ALS. Mutant PFN1 colocalizes with and enhances the formation of stress granules.\(^{167}\) As the cytoskeletal machinery is crucial in the formation and disassemblage of RNA-containing stress granules, it is tempting to speculate that mutant PFN1 may play a pivotal role in the impairment of this process. Furthermore, PFN1 mutants form insoluble and ubiquitinated aggregates, which are enhanced by ubiquitin–proteasome impairment.\(^{164}\) These observations further emphasize that the artificially separated pathogenic mechanisms in ALS are highly interrelated.

There is additional evidence for a role of axonal factors in the mechanism of motor neuron degeneration.\(^{154}\) EphA4 expression has been found to inversely correlate with disease onset and survival.\(^{166}\) EphA4 is a receptor in the ephrin axonal repellent system\(^{169}\) and induces cytoskeleton remodeling through RhoA GTPase.\(^{170}\) Its expression in ALS patients was studied because its pharmacological and genetic inhibition
rescued the ALS phenotype in zebrafish and increased the survival in ALS mice and rats.\textsuperscript{168} It was suggested that EphA4 is a determinant of the re-innervating capacity of motor neurons and contributes to the differential vulnerability of these neurons in ALS. Interestingly, ephrin signaling had already been found to be involved in the pathogenic mechanism of mutant vesicle-associated membrane protein-associated protein B and C (VAPB),\textsuperscript{171} which is a rare cause of ALS.\textsuperscript{172} Of notice, the \textit{alsin} gene, in which loss-of-function mutations cause an unusual form of ALS,\textsuperscript{173} belongs to the RhoA GTPase family.\textsuperscript{174} These observations suggest that the dynamics of the cytoskeleton and axonal outgrowth may be involved in constituting ALS and its phenotype.

More direct evidence, however, comes from the finding that \textit{tubulin alpha 4A} (\textit{TUBA4A}) mutations are associated with ALS.\textsuperscript{175} TUBA4A mutants disrupt microtubule dynamics and stability, strengthening the hypothesis that alterations of cytoskeletal integrity may have a major role in ALS.\textsuperscript{175} Interestingly, at least one mutant (W407X) displayed aggregation propensities reminiscent of other ALS-associated mutant proteins.\textsuperscript{175} Its pathogenic effect may thus be explained in terms of overburdening of the UPS and trapping of tubulin-binding proteins.\textsuperscript{175} However, the relevance of these aggregates in vivo still remains to be seen.

**A genetic basis for SALS and phenotypic variability: a complex puzzle with missing pieces**

Although unraveling of the genetic basis of some forms of ALS has provided insights in the mechanisms underlying motor neuron degeneration, many questions remain. One major issue relates to the genetic basis of ALS/FTD that occurs in patients who have no affected family members. Some of this may be Mendelian hereditary as explained earlier. Some of it may be polygenic. Genetic variants may indeed convey susceptibility, while other factors (genetic or environmental) may be needed to establish phenotypic expression. The \textit{C9orf72} expansion mutation already points in that direction, as mentioned earlier. Twin studies, GWAS, and the recent exome and whole-genome sequencing efforts address these issues.

**Twin studies in ALS**

Twin studies are valuable tools to study the genetic and environmental contribution in the etiology of complex disorders.\textsuperscript{176} Hereby, the frequency of disease occurring in both members of the twin pair is calculated (concordance) whereafter the concordance rates of identical and dizygotic twins are compared.\textsuperscript{177} High concordance rates in monozygotic twins and much lower concordance rates in dizygotic twins imply a strong genetic influence. Equal concordance in monozygotic and dizygotic twins suggests that environmental factors are more important.\textsuperscript{177}

Twin studies have been employed to test the assumption that there is a genetic component to all ALS. One study included 171 twin pairs by combining the British MND data with data from twins identified via an ALS registry in the United Kingdom and from the National Swedish Twin Registry.\textsuperscript{26} The heritability of SALS was estimated as 0.61, strongly suggesting that the majority of SALS cases also have a genetic basis. This is however an overestimation compared to two independent studies that applied genome-wide complex trait analysis (GCTA) to unrelated ALS individuals.\textsuperscript{178,179} GCTA differs from traditional GWAS studies in such a way that the overall effect of multiple single-nucleotide polymorphisms (SNPs), which may be nonsignificant on their own but additive together, can be assessed.\textsuperscript{180} The heritability estimates of SALS ranged from 20\%–25\%, implying that about one-fifth instead of more than half of disease risk comes from mutant genes in SALS.

It appears that the methodological differences make a precise estimate of ALS heritability difficult.\textsuperscript{181} The low number of possible study subjects remains a main limitation for the power of a classical twin study. In addition, recent evidence reveals the environmental, genetic, and epigenetic complexity of twin studies.\textsuperscript{181,182}

**GWAS conveying candidate ALS genes: food for thought**

GWAS have been successful in identifying the genetic basis for a long list of human diseases. Essentially, this approach screens the genome for SNPs that occur more frequently in patients with a particular disease than in control individuals. Thousands of SNPs across the genome can be examined, making GWAS an interesting tool to identify the variations contributing to a person’s risk and modifying the disease. The most consistent result obtained in GWAS in ALS is the linkage to the chromosome 9p21 locus that later was found to contain the expanded \textit{C9orf72} gene.\textsuperscript{183} Only few of the many other GWAS results in ALS reported could be replicated in different populations.

\textit{UNC13A} was identified as a risk factor for ALS\textsuperscript{184,185} and a modifier of ALS survival in general\textsuperscript{186,187} and in \textit{C9orf72} patients.\textsuperscript{188} \textit{UNC13A} is a presynaptic protein that regulates the
release of neurotransmitters, such as glutamate, at neuromuscular synapses. Upon exocytosis, presynaptic vesicles are recruited to the membrane and primed for membrane fusion. Disturbance of this priming process as a result of altered UNC13A function could lead to impairment in neurotransmitter release and ultimately to the death of motor neurons. As UNC13A regulates glutamate release, this mechanism may support the glutamate excitotoxicity hypothesis in ALS.

Polymorphisms in the elongator acetyltransferase complex subunit 3 (Elp3) gene have been found to be associated with ALS, and with survival after onset in C9orf72 expansion carriers. Elp3 is a member of a complex involved in RNA processing through mRNA elongation, histone acetylation, and modification of transfer RNA wobble nucleosides. Knockdown of Elp3 resulted in shortening and abnormal branching of motor neuron axons in zebrafish and in synaptic abnormalities in Drosophila. The effect of the polymorphisms on Elp3 function is unknown, and how Elp3 affects neuronal function is uncertain. In Drosophila, Elp3 acetylates Bruchpilot, a synaptic active zone protein involved in synaptic vesicle release. This may link its mechanism to that of UNC13A, discussed earlier. However, it is unknown whether Elp3 also affects the function of active zone proteins in mammalian cells. It is fairly well possible, given the significance of RNA metabolism as a factor in the mechanism of ALS, that it is the effect of Elp3 on transfer RNA function that explains its effect in ALS.

Disappointingly, many of the GWAS hits in ALS did not stand the test of replication. Several factors may be invoked to explain this. GWAS requires thousands of patient and control samples in order to obtain sufficient statistical power. Recently generated large DNA banks and the public availability of results obtained in individual studies, allowing meta-analyses, are necessary for this. Furthermore, in view of the disease heterogeneity explained earlier, it may be necessary to study phenotypically or genetically more homogeneous patient populations. Finally, it should be mentioned that the commonly studied SNPs only cover the most common genetic variation and that we may have to step away from this common disease/common variant hypothesis in ALS. To that end, new genotyping platforms are currently used to detect rare variants (minor allele frequency <5% instead of >5% for common variants).

Whole-genome sequencing projects and exome sequencing efforts

It is unlikely that one single locus drives the genetic risk for so-called SALS; a number of different genes may add to the risk, each with a relatively modest contribution. Exome and whole-genome sequencing covers most of all variations possible.

The use of exome sequencing, which aims to identify coding variants, led to the discovery of a missense mutation in, for example, the VCP gene that segregated with ALS and in the CHCHD10 gene that causes ALS and FTD. Exome sequencing in parent-case offspring trios has already been employed for the investigation of de novo mutations in SALS. To this end, de novo mutations have been found in genes encoding chromatin regulators, including the neuronal chromatin remodeling complex (nBAF) component SS18L1 (also known as CREST). Furthermore, the usage of exome-wide rare variant burden analysis recently resulted in the discovery of TUBA4A mutations in ALS.

Whole-genome sequencing represents a promising approach for the detection of rare variants that establish ALS risk. To this end, the “whole-genome sequencing project” aims to sequence over 1,500 genomes, making use of samples available in the UK MND DNA bank. This study collaborates with Project MinE, an international initiative that aims to sequence 15,000 motor neuron disease genomes in several participating countries (including the Netherlands, Belgium, France, the US, and Australia). The detailed clinical information linked to these samples will allow the search for genetic factors that are associated with certain phenotypes such as slow disease progression or extended survival, cognitive involvement, etc.

Disease heterogeneity and disease modifiers: hope or hype?

The finding that the same mutation can be associated with quite variable phenotypes, as mentioned earlier, suggests the existence of (genetic and/or environmental) factors that modify the phenotype. Such factors are important to identify, as they may represent therapeutic targets, which would enable us to tackle disease severity, even without knowing the exact cause of ALS. Genetic factors such as UNC13A, Elp3, and EphA4 have been mentioned earlier. In addition, effects of PGRN, PPARGC1A, APOE, MAO-B, KIFAP3, ZNF512B, VEGF, and ATXN2 on the age of onset or survival have been reported.

The intermediate-length CAG expansion in ATXN2 (≥34 repeats, encoding polyglutamine [polyQ]) was initially identified as a cause of spinocerebellar ataxia type 2 (SCA2). Afterward, the gene was put forward as a modifier of mutant TDP-43 toxicity in yeast. Based on this observation, a genetic study was performed in humans, which showed that a polyQ intermediate expansion (27–33 repeats) on one allele in...
ATRXN2 is a risk factor for ALS. This raises the possibility that SCA2 and ALS represent opposite ends of a clinical spectrum, further underlining that ALS most likely consists of a group of diseases that has motor neuron degeneration in common. In ALS, the expanded repeat alleles are interrupted with 1–3 CAA codons that reside within the CAG repeat. The number of interruptions is a determinant of disease onset. In addition, patients bearing ≥31 polyQ ATXN2 repeats had a shorter survival than those with <31 repeats. Disease-modifying therapies targeting ATXN2 may therefore represent a promising therapeutic approach for ALS.

Finally, a possible modifying effect of VEGF in ALS has been reported. VEGF overexpression or intracerebrovascular administration in SOD1-based rodent models attenuated onset of paralysis, improved motor performance, and prolonged survival. A clinical trial is ongoing in which VEGFa is delivered intracerebroventricularely using an infusion pump. The clinical significance of all these modifying factors identified in genetic studies or in animal models needs further study. An overview of genes for which there is more than anecdotal evidence is represented in Table 1 and Figure 2.

Therapeutic targets and potential for personalized treatment?

It is hoped for that the discovery of disease-causing and modifying genes in ALS allows the development of tailored therapies based upon the patient's genetic fingerprint. Approaches for this are emerging. Causal treatments directly target the factor considered to be the cause of ALS and aim to lower the expression of disease-causing proteins or RNA species. An immunological approach is being explored for mutant SOD1-associated ALS. Both passive and active immunization protocols have been shown to have a positive impact on disease progression and to delay the mortality in ALS mice by decreasing the abundance of misfolded SOD1. Genetic approaches use lentiviral vector-delivered siRNA or use intrathecal delivery of antisense oligonucleotides. They have been successfully studied in rodent models for mutant SOD1-associated ALS. Based on these findings, a Phase 1 trial was initiated and the intrathecal administration of antisense oligonucleotide against SOD1 was found to be safe and well tolerated. This study may pave the way for applying this approach to other forms of genetically determined motor neuron disorders and in fact several neurodegenerative diseases such as Huntington's disease.

The generation of iPSC-derived neurons from ALS patients now allows to generate and test novel therapeutic approaches. Some studies report these cells to show a phenotype that may be relevant to what happens in vivo, but more work is needed. One such highly interesting example is the use of neurons from patients with a C9orf72 expansion mutation. These were found to develop RNA foci reminiscent of what is seen in vivo. Antisense oligonucleotides targeting the C9orf72 transcript reduced the formation of such RNA foci suggesting that this is an approach to be considered in C9orf72 patients.

Conclusion

ALS genetics has made remarkable progress and is essential to gain insight into the complex pathogenic mechanisms leading to motor neuron degeneration and the clinical entity we call ALS. Still, despite the boom of newly found ALS genes, the cause of approximately a third of FALS and most of SALS remains unknown. The identification of the remaining genes represents a substantial challenge, for which novel approaches are being explored. The genetic heterogeneity of ALS has led to the notion that therapeutic interventions will need to be personalized. Treatment options based on the genetic cause of ALS are on the way. In the long run, further efforts will presumably enable clinicians to treat ALS patients with the most suited therapy, based on their genetic fingerprint.

Acknowledgments

The authors receive funding from the Fund for Scientific Research Flanders (FWO), the University of Leuven (GOA/11/014), the Interuniversity Attraction Poles Programme of the Belgian Federal Science Policy Office (P7/16), the European Community’s Health Seventh Framework Programme, the Association Belge contre les Maladies neuro-Musculaires (ABMM), the ALS Therapy Alliance, the ALS Association (ALSA; 039CUK and 14-LGCA-181) and the European Research Council, and the European’s Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement number 340429. WR is supported through the E von Behring Chair for Neuromuscular and Neurodegenerative Disorders, the Laevers Fund for ALS Research, and the Een Hart voor ALS and Opening the Future funds of the University of Leuven.

Disclosure

The authors report no conflicts of interest in this work.

References


145. Pearson CE. Repeat associated non-ATG translation initiation: one DNA, two transcripts, seven reading frames, potentially nine toxic DNA, two transcripts, seven reading frames, potentially nine toxic.

146. Ash PE, Bieniek KF, Gendron TF, et al. Unconventional translation initiation: one DNA, two transcripts, seven reading frames, potentially nine toxic.

147. Pearson CE. Repeat associated non-ATG translation initiation: one DNA, two transcripts, seven reading frames, potentially nine toxic.


149. Pearson CE. Repeat associated non-ATG translation initiation: one DNA, two transcripts, seven reading frames, potentially nine toxic.


151. Pearson CE. Repeat associated non-ATG translation initiation: one DNA, two transcripts, seven reading frames, potentially nine toxic.


153. Pearson CE. Repeat associated non-ATG translation initiation: one DNA, two transcripts, seven reading frames, potentially nine toxic.


155. Pearson CE. Repeat associated non-ATG translation initiation: one DNA, two transcripts, seven reading frames, potentially nine toxic.


